

DIGITAL MEDIA

A Seybold Report

Inside

2 I/O

Bob Metcalfe, the inventor of Ethernet, looks at 3DTV and wonders if networks will be the bottleneck in digital media.

12 Apple's new consumer strategy sounds familiar

Through expertise and partnerships, Apple tries to bridge the computer and consumer electronics worlds.

14 Winter CES is short on thrills . . .

. . . but digital technology is poised for a big splash with Apple's new PDAs, DCC, SkyPix and more.

15 Digital Media moves office

17 FCC allocates spectrum to interactive

Commission changes the rules of the game for local applicants.

17 'TV Answer' begs the question

Is it a case of market focus or limited vision?

19 Correction

ICOM not in video alliance.

20 Mediascape

This month, we detail video compression — from MPEG to DVI to QuickTime.

24 Events

Electronic Publishing on the 'Net' 3

Today, vast stores of information are freely available through a worldwide series of interconnected digital networks called the Internet. The Internet paradigm is increasingly being eyed by information publishers, telephone companies and computer vendors as a powerful platform for a new form of electronic publishing, where the network is more than a means to move data from the desktop to a printer. Instead, the network is where the information resides.

But what happens when information — the contents of a public library, for example, most of which is protected by copyright laws designed for print — is digitized and posted on public networks? At this point, not much. And what we do about electronic copyrights today will have great effect on the electronic publishing market of the future.

Where There's a Will, There's a WAIS 5

In 1989, Thinking Machines Corp. — the company that builds the Connection Machine supercomputer — set Brewster Kahle loose on the problem of catalyzing a market for the electronic distribution of information.

First and foremost, the infrastructure had to allow people to make money at electronic publishing. Second, it needed to be elastic enough to support anything from personal computers to consumer electronics devices to supercomputers. And third, it needed to be completely accessible and autonomous — i.e., there was to be no single point of control.

Kahle's project, dubbed WAIS (pronounced "ways"), for Wide Area Information Servers, is already well on its way to achieving those goals.

The Reading Room: A Model for Access 7

In this provocative essay, WAIS project director Brewster Kahle offers his views on the clash between private ownership of information, the roundabout of the so-called Information Age, and public access to information, which has been the cornerstone of democracy. He suggests we investigate a new model for public libraries called the Reading Room.

Where There's a Will, There's a WAIS

Public domain system solves many sticky info retrieval problems

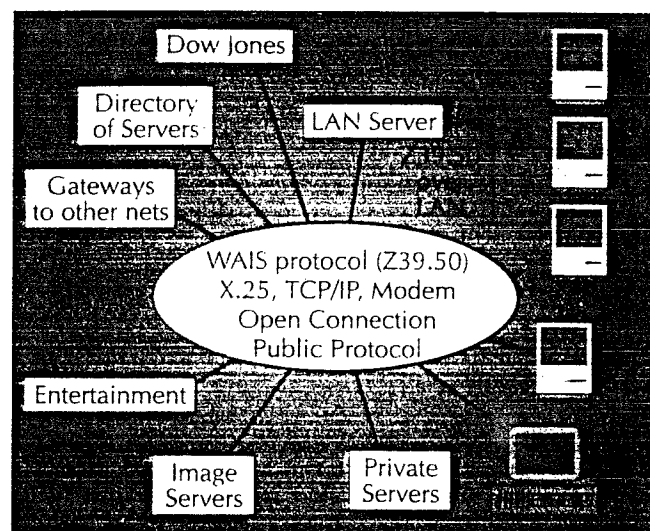
In 1989, Thinking Machines Corp. (Cambridge, MA) — the company that builds the Connection Machine supercomputer — set Brewster Kahle loose on the problem of catalyzing a market for the electronic distribution of information. Finding product (getting people to put the information online) was only part of it. The other part, much more complex, hinged on making sure that customers who wanted product could find it.

So Kahle, who's been with the firm since it was founded in 1983 and is the architect of the CPU of the Connection Machine Model 2, set out to build a system that could navigate the entire panoply of available online data sources, whether on the company's own local area network or on a Unix server halfway around the world.

The project had several goals. First and foremost, the infrastructure had to allow people to make money at electronic publishing. Second, it needed to be elastic enough to support anything from personal computers to consumer electronics devices to supercomputers, at speeds from 1,200 baud to gigabits per second. And third, it needed to be completely accessible and autonomous — *i.e.*, there was to be no single point of control. Anyone who wanted to could snap a server into the network or search a directory of servers for information.

Kahle's project, dubbed WAIS (pronounced "ways"), for Wide Area Information Servers, is already well on its way to achieving those goals.

The electronic library protocol. Even though he says he "changed it completely," Kahle built WAIS around an existing international standard called Z39.50. Once, it simply defined the client-server relationship for a remote bibliographic retrieval system. Kahle's modifications included adapting it for use on global distributed networks and adding multimedia and large document capabilities. (To be searchable, of course, data other than text has to be tagged with text.)



The WAIS Architecture

The modified Z39.50 protocol is now endorsed by the Library of Congress, Apple, Sun Microsystems, Dow Jones and Mead Data Central.

Unhinging client and server. What's unique and powerful about the WAIS protocol is that it has unhinged the connection between the client computer's user interface to the system, where a user like me originates a request for information, and the server that translates and acts upon it.

This facilitates a number of important things. As a user, all I have to learn in order to look for information on the network is the WAIS user interface. I don't have to learn to navigate the Internet, or Nexis or Lexis, or CompuServe, or even my own corporate SQL database. If a server is registered on the WAIS system, all I have to do is type — in English — a few words about what I'm looking for in the WAIS text field labeled, "Look for documents about."

The client computer encodes my request in "WAIS-speak" and sends it out to the worldwide network of WAIS servers. The servers translate the query (I don't need to know how), find articles they think match my request, and send them back to me.

I look at the "hits" and select those that look most applicable. Based on that feedback, I can send the query out again until I've gotten the specific information I was looking for. (The method is called "relevance feedback," and it's proven to be a very efficient way to hone in on information.)

Charging for information. This "decoupling" of the client from the server is also a powerful tool for electronic

publishers who want to charge for their information in a variety of ways. Control of the charging structure, says Kahle, is resident at the server.

Though how that will evolve isn't clear yet, he says, the technology doesn't stand in the way of any method that publishers might want to use, including subscriptions, document transfer fees, or any combination they can think of. If a piece of information I'm looking for happens to be available for a fee, the server will send a message back to me telling me so and maybe asking me to take some action authorizing the purchase if I haven't already tacitly done so.

Keeping within the law. Another of the system's many useful features (which are far too numerous to list here) is its ability to create "document pointers." These pointers, which stand alone from the document itself, note where a piece of information is located in the network without making a copy of the document itself.

Thus, I can store a pointer so I can return to a document easily, or pass along a pointer to someone who can then find that document. Neither of us is violating copyright laws because we haven't copied or passed along an actual document to someone else. It's the electronic version of the International Standard Book Number, or ISBN, that today makes it possible to locate books in the physical universe.

I can also set up what's called a "dynamic folder" for ongoing topics of research or investigation, whereby WAIS will constantly or periodically (it's up to me) update the folder with new material. All in all, a rather nifty pressure nozzle for the information firehose.

Gating issues

The reason why Kahle's company would bother going to all this trouble is that it believes the Connection Machine is the very best database server for this nascent electronic publishing market (of course).

But unlike what some companies have and might have done under the same circumstances, Thinking Machines has placed the WAIS source code in the public domain to promote widespread use of the protocol no matter what brand the server.

Kahle realizes the danger of trying to set a standard so early in the market development cycle. But, he says, "I'm working with standards committees to make sure that what we're doing is not a proprietary system, but also that it's free to evolve."

The value of the WAIS system is evident by how quickly new servers are connecting. As of this writing,

145 Internet servers were connected via the WAIS protocol; when we spoke with Kahle less than a month earlier, the number was 120.

A Unix interface already exists, as do VT100, Macintosh and NeXT versions. DOS and Windows are on the way. The Library of Congress, which boasts 25 terabytes of data, has plans to make its catalog available via the protocol.

No "for pay" servers. Despite Thinking Machines' desire to jump-start an electronic publishing market with WAIS, there are still no external interfaces to "for pay" services published in the public domain, though some are under development. Such published interfaces are vital if the WAIS system is to be useful as a real publishing system, not just as a nifty new trick for hackers to play with.

However, progress seems destined to proceed apace, as Kahle and other WAIS supporters gathered at Research Triangle Park in North Carolina on February 3 and 4 to launch the North Carolina WAIS Initiative consortium. The consortium will encourage broad-based development of WAIS interfaces and services and support WAIS freeware.

"The New York Times Review of Servers"?

The potential of the WAIS architecture lends itself directly to the question of what kinds of information might start pouring onto the net. Many established companies, including Dow Jones and Apple, are working closely with Kahle. Dow Jones, in fact, is putting up a test "for pay" WAIS server on its DowVision network with the *Wall Street Journal*, *Barron's* and 450 magazines.

Today, published WAIS servers include a directory of servers, the CIA World Factbook, a partial patent database, databases on molecular biology, a poetry server at MIT, cookbooks, descriptions of government software packages, and weather maps and forecasts.

The potential for the number of published WAIS servers to mushroom in number even caused *Byte* magazine to speculate that an independent agency such as *Consumer Reports* might create a rating service to monitor and rate servers in the WAIS directories. The idea would be to serve as an independent guide to quality, publicly tagging servers that regularly delivered bad information or didn't work at all.

What a different world that would be!

Denise Caruso 